Paper

Ontology Creation Process in Knowledge Management Support System for a Research Institute

Cezary Chudzian

Abstract—Though the issue of knowledge management is a hot subject of interest in nowadays market companies, integrated solutions fit to the specific needs of research institutes still require more attention. This paper documents a part of the research activities performed at National Institute of Telecommunications, related to development of research institute knowledge management support system. The ideas lying in the background of the system come from the recent theories of knowledge creation and creativity support and from experience with everyday practice of knowledge management in market companies. Main focus is put here on the issue of creation of a research topics ontology that is meant to be a semantic backbone of the system. Three-stage approach is proposed, aiming at the construction of ontologies for different levels of organizational hierarchy, from individual researcher, through group or unit, up to the whole institute. Created ontologies are linked to knowledge resources and support diverse activities performed at those levels.

Keywords*— *creativity support, knowledge creation, ontological engineering, scientific knowledge management.

1. Introduction

Knowledge management has become recently a hot topic not only in many research communities all over the world but also increasing interest may be noticed in market companies getting lost in their own information sources of different kind. At the same time arising concept of Business Intelligence 2.0 is moving towards proactive approach to problem solving in business environments, instead of reactive, latter being implemented as searching for patterns in collected information to improve future decisions (knowledge discovery and data mining). In order to act before an event occurs, one must have a rapid access to the sources of knowledge critical to his decisions. Not only written documents, but also multimedia content [1] and domain experts are the subjects of knowledge management as important explicit and tacit knowledge resources. Our experiences with work in a big telecommunicatins market company, shows that the issues of knowledge management are in a very immature stage in there, but at the same time there is quite a big need to implement some adequate solutions in the area.

Even more complicated sounds a question of how to manage knowledge in a research company or institute, where problems being solved are usually more complicated than those in the commercial environments. Moreover, the final product of a research institute is the knowledge itself. Thus there is a strong need to organize and support development of creative environments [2] to improve the quality of research.

Ontologies as a knowledge representation method became very popular[1] in recent years, especially in computer science community after popularizing the idea of semantic web [3] as the future of the Internet. Number of standards, tools and languages supporting the idea has been developed since then. But on the other hand, it is hard to name a set of mature, well-established and widely used software implementations of knowledge management for market or research use[2].

Ambition of our research group in National Institute of Telecommunications (NIT) is to create the research institute knowledge management system (RIKMS) fit to the needs of a research institute, based on ontology engineering tools and methods and employing the ideas of creative space described in [2], [4].

This preliminary paper is mainly devoted to the problems of generation and maintenance of research topic ontology and is structured as follows. Section 2 discusses differences between two main ontological views on activities of a research institute. Some remarks on the way an ontology of research topics shall be structured are presented in Section 3. Section 4 is devoted to the general framework proposed for ontology creation and maintenance while Section 5 summarizes the paper and gives some directions of future development.

2. Organizational and Topic Ontology

The RIKMS reflects two different, but interrelated, perspectives on the knowledge maintained in a institution. The former is concentrated on the organizational aspects or on how to organize knowledge intensive processes and the latter is focused on research topics lying in the field of interest of the institution, or on how to refer to research areas and topics.

Building blocks of the organizational ontology reflect the structure of an organization, its working regime, accepted standards, policies and procedures, worked out prod-

[1]At least judging from the number of books and publications in this area.

[2]Contrarily to, e.g., knowledge discovery and data mining tools.

ucts, etc. Organizational ontology is usually hardwired in a software system dealing with its concepts, taking form of relational tables with fixed structure and hard-coded procedures expressed in some chosen language. On the other pole would be a software system storing and processing organizational ontologies in more general form as constructs of one of the ontology engineering languages[3]. However this is a completely different subject lying rather in the scope of software engineering methods and will not be discussed here.

By topic ontology we mean here a set of interrelated topics researched by the institute. It creates a different point of view on the activities performed within the institute. Organizational ontology may be seen as orthogonal to topic one and consequently they intersect each other. Projects may be indexed with keywords taken from the topic ontology of the institute, similarly employees will manifest competences in some topic labelled research areas.

Research and scientific institutions usually tend to form a hierarchical organizational structure constituted of departments, divisions, laboratories or working groups, centered around some research issues. Every unit of an institute has its own leader and employs people with similar educational and scientific background. Furthermore intellectual heritage and common sense of every group of this kind has been formed by its history, tradition, shared values, cooperation with external partners, long-term project experience. Such a group is thus thematically homogeneous to some extend and hermeneutic horizons[4] of its members are more coherent with respect to one another than to members of the other organizational units.

Reverse influence of research topics on organizational structure may be observed as well when we consider the origins of units within an institute. They are often formed around charismatic leader, transformed from successful working groups or answer a need to undertake a research in some previously uncovered area.

Summing up organizational and topic ontologies are closely bounded and one cannot drop any of them when dealing with the subject of knowledge management in a research institute. Our approach to ontology construction utilizes organizational structure as a framework for topic ontology creation and maintenance.

3. Topic Ontology Representation for a Research Institute

There are several motivations for creating a topic ontology for a research institute. Ability of viewing processes and their outcomes from the perspectives of projects, their products or people involved in them is attainable as all they are distinguishable concepts of organizational ontology and thus might be somehow reflected in the structure of knowledge management software system. But the questions immediately arise of how an overview of the activities from the research topics perspectives may be achieved or what the set of all topics researched by the institute consists of. Possibility of taking topic centered perspectives on projects, products, employees and documents has a meaningful importance for people involved in management of a research institute and heads of its departments. It supports many decision making tasks. Lets enumerate some of them.

- Reporting the achievements in particular fields of scientific activity entails a reflection on appropriateness and up-to-dateness of current organizational structure and enables build-up of development strategies.
- When applying for a new project or analyzing research trends, topic ontology centered view helps to determine whether an institute has enough expertize in related thematic fields.
- Knowing the competences of individual employees is a key prerequisite for building up interdisciplinary working groups capable of dealing with complex problem with many diverse research threads.
- Analysis of structure of topic ontology may lead to identifying the germs of new research topics.

On the other hand, topic ontologies may be useful at the individual researcher level as the important input for the tools supporting creativity. The idea of hermeneutic EAIR (enlightenment, analysis, hermeneutic immersion, reflection) spiral of searching through rational heritage of humanity and reflecting on the object of study has been presented in [4] with experiments on ontology supported hermeneutic agent, helping a user in search for knowledge sources related to object under research, reported in [2]. The ontology is used there to define researcher semantic profile that machine is able to process and use in order to help in finding relevant knowledge resources on the world wide web.

Textual and multimedia information is not the only source of knowledge in the research institute. Having an access to semantic profiles of institute employees, software agent might locate a person with strong competences in the subject of study. It could be a hard task for someone not familiar with everyone's research interests, but a computer fed with profiles of individuals can be very helpful.

3.1. Local versus Global Ontology

Different applications of topic ontology demand different views on set of concepts and relations. Intuitively, at the individual level, granularity is to be greater and ontology more detailed, as it supports actions performed during everyday work, at rather operational level. Higher in the or-

[3]Many standards have been developed. Let us mention web ontology language (OWL), or lower-level resource description framework (RDF).

[4]Hermeneutic horizon following H.-G. Gadamer [5] is "*The totality of all that can be realized or thought about by a person at a given time in history and in a particular culture*".

ganizational hierarchy of the institute, more general views on topics are needed as the horizon of group activities is more strategic.

Not every concept and relation is meant to be visible at the higher levels, some of them may remain private, but those of higher levels must be more reliable, commonly agreed and formal.

Distinctions mentioned above, along with general remarks on orthogonality of organizational and topic ontologies (see Section 2) lead to conclusion that it seems to be more reasonable to maintain distributed ontologies associated with different levels of organizational hierarchy, from individual, through group, ending up with an overall ontology of the institute.

As the responsibility for communication of an organizational unit with its environment lays on unit or, more generally, group leader, his or her role in ontology construction and maintenance processes should be superior. The leader is to be especially involved in the mechanism of keeping ontology of his group consistent and integrated with those of higher levels. We shall emphasize the role of the leader in proposed framework.

What must be decided next is whether there should be one ontology defined globally for the whole domain[5] and then adapted by its constitutive units or the better solution is rather to develop local ontologies for all individuals and combine them into higher level ontologies. We believe that the bottom-up strategy is a better solution. The intuitive justifying argument is that in the top-down procedure, there must be overall ontology defined, detailed enough to help in creativity support processes and, at the same time, covering all possible topics lying in the field of interest of the institute. Lets assume, we wish to adapt some kind of telecommunications ontology defined by ITU[6]. Saying that, e.g., National Institute of Telecommunications covers the whole universe of telecommunication related issues as defined by ITU, and nothing more, is not neccessarily a true statement[7]. From the one point of view that would be a nice property as it could enable adaptation of single, global view on activities of all telecommunication institutes all over the world. But at the same time it introduces informational mess, by importing to institute's research field concepts that are out of its scope and forgetting those which are applicable. The question of topic map for the whole institute would remain unanswered.

Attemps to build up a NIT ontology ([6] and further work), showed that the institute is active in a variety of diverse and advanced research fields, including typical low and high level telecommunications problems like electromagnetic compatibility, radiocommunications and mobile telephony, optoelectronics, network infrastructure and management, but also law, social and market issues like regulatory problems, customer satisfaction surveys and decision support problems including knowledge discovery and management, game theory and logic.

Having above considerations in mind, it seems to be much more promising approach to build an ontology starting from individual level, promote their local concepts to higher levels of department and institute in some manner and integrate them to achieve the global picture of institute activities. Such an idea of heterogeneous ontologies in distributed environment has been discussed in [7]–[9].

We stated above that every individual and group hold their own ontology. Group has been defined as an organizational unit, like department or laboratory. However within an institute there may exist a number of task teams, interdisciplinary groups with people primarily affiliated with different organizational units. Research institute as the whole may be seen as a group too. All those meanings of a group should be enabled to have their own common ontologies. In such a context, group leader is a head of an unit, or director of the institute, but also informal group leader or a person designated to take care of public image of, and knowledge management in his or her group.

3.2. Light versus Heavyweight Formalism

The word ontology has been being present in the common vocabulary for several centuries. Following is the Wikipedia's definition of its philosophical connotations:

> ... the oldest extant record of the word itself is the Latin form ontologia, which appeared in 1606, in the work Ogdoas Scholastica by Jacob Lorhard (Lorhardus) and in 1613 in the Lexicon philosophicum by Rudolf Göckel (Goclenius). The first occurrence in English of "ontology" as recorded by the OED appears in Bailey's dictionary of 1721, which defines ontology as "an Account of being in the Abstract".

Next search on the word ontology in Wikipedia, but this time in the context of computer science, gives next definition:

> ... an ontology is a formal representation of a set of concepts within a domain and the relationships between those concepts.

The key word above is "formal". Increasing the amount of formalism gives many different realizations of ontology. Classification framework has been proposed in [10], while referring paper [11] discusses, also interesting from our point of view, issues of ontology based documents annotation. Lets enumerate definitions taken from [10] in a light-heavy order:

controlled vocabulary – finite list of terms,

glossary – list of terms and their meanings,

thesauri – list of terms with synonym relationship defined (but not an explicit hierarchy),

is-a – hierarchy of classes and their instances,

[5] Here: research institute.

[6] International Telecommunication Union – United Nations agency for information and communication technologies.

[7] Where is the place for this paper in such a case?

frames – classes having properties (including relations to the other classes),

value restrictions – characteristics of properties (e.g., type/class, value) restricted,

disjointness, inverse, part-of – additional relations between classes with well-defined semantics,

general logical constraints – arbitrary logical statements on classes, properties and instances.

Most of people require at least "is-a" level definition to be achieved to consider specification of the domain to be an ontology.
In our case it is still unclear how formal the model will be. We initially assume it to be at "is-a" level with possibility of defining additional, other than "is-a", somehow restricted, relations between classes. Nevertheless this setup may be significantly modified while the system evolves.

4. Topic Ontology Creation and Maintenance Process in a Research Institute

In this section a general framework of the process of ontology creation and maintenance for a research institute is presented. Must be stated that the work documented below is at the early stage and only limited number of details may be provided apart from the general idea.
In the following, for the sake of brevity, a set of interrelated ontological concepts associated with person, unit or institute will be called an ontological or semantic profile or simply a profile.
The basic scenario for ontology construction starts with submission of the new knowledge resource to the RIKMS. It is then analyzed by some automatic concept extraction method from text[8].
Then phase of individual[9] reflection is initialized, with querying an user on a relevance of discovered topics to his or her individual profile, their relations to existing topics in the profile and profiles of other people or higher level profiles and thus stimulate user to take a reflection on the profile and modify it accordingly. This step may be viewed as a limited form of analysis transition in EAIR spiral [4] as a localization of the new concepts in the context of existing semantic profiles.
Submitted document will be then indexed with new topics signatures and may serve as a proof of user's competence in the field characterized by those topics. Moreover, if submission occurred in some specific context, as a final report from project or part of some other activity, then system stores the relation of topics, document and context for further use.

[8]At the moment we assume knowledge resources to be textual documents.

[9]Local or in other words happening between software system and its user.

Final cross-level agreement stage starts with identifying those parts of ontology which could be promoted to one of the higher levels of organizational hierarchy. During the process of system guided debate between all interested parties, concepts and relations are delegated up the hierarchy on the basis of common agreement, but with the deciding vote of a group leader. Cross-level agreement stage in the context of knowledge creation theories may be seen as a counterpart of debating transition of EDIS (enlightenment, debate, immersion, selection) spiral [4].

4.1. Topic Ontology Generation

This section describes document-based concept extraction method, being automatic step of ontology construction process. Different methods of automatic ontology construction have been surveyed in [12]–[14]. The approach proposed in this paper is similar to those used in On-To-Knowledge project for retrieval of relations between concepts from documents [15], [16].
In [6] a number of tools for automatic concept extraction have been outlined and the results of experiments conducted on publications from *International Journal for Telecommunications and Information Society* have been presented. Special attention has been paid to OntoGen system [17]. The reader is encouraged to refer to [6] for more details.

OntoGen generates an ontology in a semi-automatic process on the basis of a corpus of documents describing specific domain or subdomain of interest. Document clustering based method is used which needs some minimal number of documents to be available. It is rather designed for attaining a global view of a domain under investigation.

Chosen for our framework topic generation routine utilizes the well-known idea of frequent itemsets discovery in transactional data [18].
In preliminary step document is being transformed into transactional data by decomposing it into a set of sets of words (wordsets). Each wordset is roughly equivalent to a sentence in a grammatical sense.
Frequent wordset is defined as follows. Having a set of wordsets $D = \{S_i\}_{i=1,..,N}$ frequent wordset is a set of words $S^* = \{w_k\}_{k=1,...,M}$ simultaneously contained in at least *Supp%* of wordsets from *D*, where *Supp* is called a support of a wordset:

$$S^* \text{ is frequent } \iff \frac{|\{S_i \in D \mid S_i \supseteq S^*\}|}{|D|} \geq Supp\,.$$

The motivation behind searching for frequent wordsets in a document is rather straightforward. If the same group of words has been used by authors in a number of sentences it may designate a concept or a set of related concepts.
After frequent wordsets are discovered, second operation is performed, so called pruning step. It is obvious that if a wordset S_i^* is frequent then every wordset S_j^* being a subset of S_i^* is frequent as well, as it is at least contained in the same set of sentences as S_i^* is. For the sake of brevity, all wordsets being subsets of frequent wordsets are removed

from the result in a pruning phase. All, but not those which exist in significantly greater fraction of sentences than their supersets and thus might indicate more general concept.

Algorithm 1: Frequent word sets mining for ontology construction

```
Data: document
Result: nondominated_frequent_groups
norm_document = ConvertAndPreprocess(document)
transactions = MakeTransactions(norm_document)
frequent_groups = Apriori(transactions)
nondominated_frequent_groups = Prune(frequent_groups)
```

Table 1
Example of frequent word groups

Words	Count
Paper I	
theory, rational, intuition	16
normal, creation, knowledge	14
spiral, creation, knowledge	13
creation, knowledge	51
seci, spiral	19
humanity, heritage	16
space, creative	16
civilization, knowledge	15
dimensions, creative	14
shinayakana, systems	13
approach, systems	13
tacit, knowledge	13
heritage, knowledge	13
Paper II	
triple, helix, creation, knowledge	6
processes, creation, knowledge	10
indicators, quality, reference	9
spirals, creation, knowledge	8
normal, creation, knowledge	7
academic, creation, knowledge	7
spiral, creation, knowledge	7
minimized, maximized, indicators	6
units, sets, data	6
maximized, indicators, quality	6
best, sets, data	6
sets, data	14
profile, reference	12
Paper III	
nakamori, wierzbicki, spiral	10
academic, creation, knowledge	9
technology, management, knowledge	9
science, systems, knowledge	9
science, management, knowledge	9
creation, knowledge	34
management, knowledge	27
science, knowledge	18
nanatsudaki, model	16
seci, spirala	12
processes, knowledge	12
academic, spirals	10
processes, creation, knowledge	10
Paper IV	
representation, multiple, aggregation, criteria, knowledge	13
coefficients, weighting	19
integration, knowledge	15
form, knowledge	15
sum, weighted	14
reservation, aspiration	13
aspiration, levels	13

The four steps procedure starts with the preparation of document to make it fit to input requirements of the word groups searching routine. Format conversion to plain text, font encoding translation, removal of stop words, lowercasing and stemming are the main steps in preprocessing phase. After accomplishing that part, document is transformed into transactions that are fed to a frequent wordsets discovering algorithm. Finishing pruning step reduces the number of word groups and gives the final result.

Table 1 shows frequent word groups with cardinality greater than two for some papers thematically located in the field of knowledge science.

4.2. Individual Reflection

The purpose of the next stage is to populate individual profile with newly discovered topics and to define relations to the concepts found in both individual and neighbourhood profiles. It takes form of interaction between user and software system driven by word groups discovered in previous step and current structure of profiles. Sovereignty of the user is a superior principle. He or she decides on the final shape of individual profile. The role of software system is in stimulation of user's reflection by requesting advice about a local hierarchy of ontology. It might be seen as an engine searching for new topics and relations in both new source of knowledge and already established semantic structures. The final decision is always left to the user.

System is detecting whether any of basic indicators of new concept or relation existence arise and should be reported to the user providing a new source of knowledge.

Frequent word groups. Automatic topic generation phase proposes a number of words – candidates for new topics and indicates their coincidence in contexts of sentences. The n-grams for $n > 1$ are expected to be more informative than unigrams. They may carry two meanings. One is that there exists a concept in the domain identified by the name consisting of more than a single word. Second interpretation is a set of related concepts. Mix of those two is possible as well. Such a hypothesis of distributional semantics[10] lies in the basis of, e.g., some of the text summarization systems [19]. Scenario for this step is to present frequent groups to the user and let him decide whether they can contribute to his individual semantic profile and if so then whether they form single concept or group of interrelated topics. This may be enhanced by presenting some additional information like showing the quotations from the submitted document in which they appear.

Super- and sub-wordset. If user designate word groups being in a subset-superset relation, as bases for new concepts, it may be an indicator of existence of one of important relations "is-a" or "part-of" between corresponding concepts. System should suggest such a solution.

[10]Context has a strong influence on the word meaning.

Integration with existing profiles. New topics and relations evolve in the semantic context of person and neighbourhood. Thus during questioning process, system should take into consideration existing concepts and ask the user to localize newly discovered ones in the whole semantic profile. Searching for counterparts of ontological concepts is a subject of research in the area of ontology matching and alignment [20]–[22]. Variety of techniques are available from simple matching by name ending up with more sophisticated methods. This issue still needs to be investigated in more detail.

User invention. System should be able to process any other modifications proposed by the user at this stage. However some kind of constraint for user's freedom should be applied. It might be an obligation to provide an explanation of why the modification had been made. System may ask for a reference to a source of information on the new topic as a kind of evidence or to place new elements in the current ontological structure by linking them to existing topics.

4.3. Cross-Level Agreement

As mentioned in Subsection 3.1 the framework we are aiming at shall generate global view of the topics researched in the institute by aggregating individual profiles of employees. Special role is granted to a group leader, who has a final deciding vote as a person responsible for the overall picture of activities performed by his or her group.
Software system is engaged in two aspects of creating an agreement on ontology structure. First, it again stimulates a reflection on how the higher level structure should look like, by informing individual users on existence of potentially promotable concepts and relations. A number of indicators might be considered. Below a couple of exemplary ones are listed.

Shared concept. Sharing a concept between two or more individual profiles seems to be a good reason to promote it to the higher level. Both profile holders should be notified about the match found and decide together whether publish the concept or not.

Shared relation. One meaning of relations sharing is analogous to concept sharing. The second one is that relation linking concepts that had been promoted to the higher level should get high score as well.

Superconcept. Some of the relations are distinguished among others. For instance "is-a" associating super- and subconcept play a special role in any system as it introduces a hierarchy into it. Therefore superconcept of a concept promoted to the higher level should get a high score.

Strongly supported by the sources. Concepts from a individual profile having many knowledge sources associated are more likely to be promoted.

Existing ontology. Relations and concepts imported from another, especially widely recognized ontology are desirable. However they first should appear in at least one individual profile to justify their relevance to the institute.

User invention. Again system should give its users freedom to promote concepts and relations they consider to be important.

The second task for software system within a process of achieving cross-level agreement is assistance in debating on the shape of higher level structure. After a part of individual ontology is proposed to be promoted by the user, system should notify group leader and provide him or her with all acquired information about new ontological findings. After the final decision is made system should propagate updated ontology among all parties that may be interested in it and, in case of need, initiate a debate leading to the agreement.

5. Conclusions and Future Work

In this paper the framework for ontology construction for a research institute has been proposed. The framework is organized in a distributed hierarchical structure, with local ontologies associated with individual employees and an integrated higher level group ontologies with concepts and relations promoted from individual profiles. Three main steps of ontology construction have been outlined, namely topic generation from documents, individual reflection on ontological profile and cross-level agreement between interested parties. Special, superior role of group leader has been emphasized. Some preliminary results for simple, but robust topic generation method have been presented.
We believe the framework may be a better choice for a research institute trying to develop its own ontology of research topics for integration and management of knowledge resources than adaptation of any well-known domain ontology or creation of global ontology by domain experts. Reflection and agreement stages have themselves an additional value as they are driving processes of exploring scientific neighbourhood (individual reflection) and exchanging knowledge through debate (cross-level agreement). As such they may be seen as supporting creativity in scientific environment.

What must be stressed here is that development is at the early stage and far from complete. There is still much of work to be done. More sophisticated methods for topic extraction from documents are to be tested, detailed specification of reflection and agreement phases and implementation of software component with appropriate human-computer interface is still to be worked out.

References

[1] *Semantic Multimedia and Ontologies. Theory and Applications*, Y. Kompatsiaris and P. Hobson, Eds. New York: Springer, 2008.

[2] *Creative Environments: Issues of Creativity Support for the Knowledge Civilization Age*, A. P. Wierzbicki and Y. Nakamori, Eds. *Studies in Computational Intelligence*. Berlin: Springer, 2007, vol. 59.

[3] T. Berners-Lee, *Weaving the Web: the Original Design and Ultimate Destiny of the World Wide Web by Its Inventor*. San Francisco: HarperSanFrancisco, 1999.

[4] *Creative Space: Models of Creative Processes for the Knowledge Civilization Age*, A. P. Wierzbicki and Y. Nakamori, Eds. *Studies in Computational Intelligence*. Berlin: Springer, 2006, vol. 10.

[5] H.-G. Gadamer, *Truth and Method*. 2 ed. New York: Crossroad, 1989.

[6] A. P. Wierzbicki and E. Klimasara, "Wybrane problemy zarządzania wiedzą. Zadanie 1: Praktyczne aspekty zarządzania wiedzą", Statute Project nr 06300017 Internal Report. Warsaw: National Institute of Telecommunications, 2007 (in Polish).

[7] M. Uschold, "Creating, integrating and maintaining local and global ontologies", in *Proc. First Worksh. Ontol. Learn. OL-2000 14th Eur. Conf. Artif. Intell. ECAI-2000*, Berlin, Germany, 2000.

[8] P. R. S. Visser and Z. Cui, "On accepting heterogeneous ontologies in distributed architectures", in *ECAI'98 Worksh. Appl. Ontol. Probl.-Solv. Meth.*, Brighton, UK, 1998.

[9] A. Kuczynski, D. Stokic, and U. Kirchhoff, "Set-up and maintenance of ontologies for innovation support in extended enterprises", *Int. J. Adv. Manuf. Technol.*, vol. 29, no. 3/4, pp. 398–407, 2006.

[10] O. Lassila and D. L. McGuinness, "The role of frame-based representation on the semantic web", Technical Report 01-02, Knowledge Systems Laboratory, Stanford University, 2001.

[11] O. Corcho, "Ontology based document annotation: trends and open research problems", *Int. J. Metadata Semant. Ontol.*, vol. 1, no. 1, pp. 47–57, 2006.

[12] "A survey of ontology learning methods and techniques", A. Gómez-Pérez and D. Manzano-Macho, Eds. OntoWeb Deliverable D1.5, Universidad Politécnica de Madrid, 2003.

[13] P. Buitelaar, P. Cimiano, and B. Magnini, *Ontology Learning from Text: An Overview*. *Frontiers in Artificial Intelligence and Applications Series*. Amsterdam: IOS Press, 2005, vol. 123.

[14] C. Biemann, "Ontology learning from text: a survey of methods", *LDV-Forum*, vol. 20, no. 2, pp. 75–93, 2005.

[15] J. U. Kietz, R. Volz, and A. Maedche, "A method for semi-automatic ontology acquisition from a corporate intranet", in *Proceedings EKAW-2000 Workshop "Ontologies and Text", Juan-Les-Pins, France, October 2000*, *LNAI*, vol. 1937. Berlin: Springer, 2000.

[16] J. U. Kietz, R. Volz, and A. Maedche, "Extracting a domain-specific ontology from a corporate intranet", in *Proceedings of the Fourth Conference on Computational Natural Language Learning and of the Second Learning Language in Logic Workshop, Lisbon, 2000*, C. Cardie, W. Daelemans, and C. Nédellec, Eds. New Jersey: Association for Computational Linguistics, 2000, pp. 167–175.

[17] B. Fortuna, M. Grobelnik, and D. Mladenic, "Semi-automatic data-driven ontology construction system", in *Proc. 9th Int. Multi-Conf. Inform. Soc.*, Ljubljana, Slovenia, 2006.

[18] R. Agrawal, T. Imielinski, and A. N. Swami, "Mining association rules between sets of items in large databases", in *Proc. 1993 ACM SIGMOD Int. Conf. Manage. Data*, P. Buneman and S. Jajodia, Eds. Washington, USA, 1993, pp. 207–216.

[19] C.-Y. Lin and E. H. Hovy, "The automated acquisition of topic signatures for text summarization", in *Proceedings of the 18th International Conference on Computational Linguistics, Universität des Saarlandes, Saarbrücken, Germany*. San Francisco: Morgan Kaufmann, 2000, pp. 495–501.

[20] M. Ehrig, *Ontology Alignment: Bridging the Semantic Gap (Semantic Web and Beyond)*. New York: Springer, 2006.

[21] M. Ehrig and J. Euzenat, "State fo the art on ontology alignment", Knowledge Web Deliverable D2.2.3, University of Karlsruhe, 2004.

[22] N. T. Nguyen, *Advanced Methods for Inconsistent Knowledge Management (Advanced Information and Knowledge Processing)*. New York, Secaucus: Springer, 2007.

Cezary Chudzian received his M.Sc. in computer science from the Warsaw University of Technology, Poland, in 2002. He is a researcher at the National Institute of Telecommunications. Currently he works on his Ph.D. in the area of knowledge management. His main scientific interests include: practical applications of knowledge discovery techniques, machine learning theory, knowledge management, global optimization, and advanced software engineering.

e-mail: C.Chudzian@itl.waw.pl
National Institute of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland